

Navigating through a Document-Centered Electronic Medical Record: a Mock-Up Based on WWW Technology*

Jacques Bouaud, PhD., Brigitte Séroussi, M.D., PhD.

DIAM: SIM/AP-HP & Département de Biomathématiques, Université Paris 6,
91, boulevard de l'Hôpital, F-75634 Paris Cedex 13, France
{jb,bs}@biomath.jussieu.fr

Current WWW technology facilitates the development of "hypertext" applications. A hospital-wide study of users' requirements in France led to a document-centered approach to the patient Electronic Medical Record (EMR). In order to refine such a specification, and taking advantage of WWW technology, we have developed a running mock-up of a document-based EMR from an actual paper-based patient record. Synthesis documents were added and linked to original replicated paper documents to form a hypertextual EMR. The mock-up has been presented to health care professional boards to gather their remarks and wishes, and then enhanced accordingly. The current version reflects (part of) their requirements for an EMR, and is presented in this paper.

INTRODUCTION

The World Wide Web¹ (WWW) is devoted to serve information over the internet. Its success essentially relies on the shared normalization of resource exchange through an exchange protocol (HTTP), a name space for resources (URL), and a mark-up language for structuring documents (HTML). Remote client "browsers" interpret HTML documents delivered by information servers to display them. These documents basically contain text and images, and particularly links to other identified documents. All the documents with their links constitute a huge hypertext. The design of a hypertext is not an easy task since one has to determine the granularity level of documents with their inter relationships in order to facilitate information gathering by readers while preventing the risk of their getting lost. Nevertheless, this normalization provides a cheap and easy-to-use technology to develop arbitrary client-server applications over heterogeneous computer networks that is today heavily exploited.

In the medical domain, many real and practical WWW applications have been developed. Many of them focus on accessing existing databases, typically medical/educational repositories^{2,3,4} or even patient clinical data^{5,6}. Consequently, the proposed "hypertext" model replicates the underlying database model, and therefore constrains information retrieval capabilities.

*This work has benefited from the financial support of the "Telematics for Language Engineering" programme of the European Commission (project DOME, MLAP #63-221).

Within a European project, we have carried out a study to design a document handling system for use in the health care environment. The originality of the DOME survey conducted at "La Pitié-Salpêtrière" hospital (Paris, France) has been to design a target system which is not technology driven, but which addresses the global end-users' needs for computational aids to produce and process medical. Health care professionals were involved in the specification of an ideal system which is best described from a user point of view as a multimedia, hypertextual document-based Electronic Medical Record (EMR) with browsing functionalities.

In order to refine the initial specification, we developed a running mock-up of a document-based EMR using WWW technology from actual medical documents of a Paper-based Patient Record (PPR). We were mostly concerned with hypertextuality needs in such an approach: basically, which (new) documents are needed, how they are related, how they should be accessed, and, in a lesser extent, how they should be presented. The mock-up was presented several times to health care professional boards and enhanced according to their remarks and wishes. This paper describes the current state of the mock-up which reflects (part of) their requirements for an EMR.

BACKGROUND

The Electronic Medical Record

A central element in medical informatics developments is the EMR and its advent in daily practice. EMR proposals usually follow two trends which, in the extreme, consist either in the replication of the PPR or in the structuring of patient related data. In the future, both aspects should be mixed.

When EMRs tend to be highly structured⁷ they enable sophisticated exploitation of medical data. But this approach had been considered to be too restrictive, in terms of medical expressiveness, and fails to gain a wide consensus for communication. The current practice of health care actors mainly relies on free text reports and notes^{8,9}. Besides, another challenge for the EMR is to provide easy and consistent access to patient information which is comprehensive, credible, accurate and relevant, and to provide this information at any time and at the place of care, according to pa-

tient confidentiality standards. In this respect, the PPR, though imperfect, exhibits in practice some properties that facilitate physicians' reading^{10,11}. It seems desirable that such features are taken into account when designing user interactions for the EMR to be accepted¹².

Users' Requirement Survey

Context. La Pitié-Salpêtrière is one of the largest hospitals in Europe, with 62 departments, 7 000 employees, 2 430 beds. In this huge institution, a market survey has been conducted to assess the present situation in terms of medical reports production and processing, and to elicit clinicians' expectations for an ideal document processing system related to the EMR.

The survey has been carried out by means of personal interviews and document analysis on 36 departments. Around 150 persons including both doctors (residents, senior residents and heads of departments) and secretaries have been interviewed about their current practice in handling patient medical records.

Recommendations. French clinicians operating at La Pitié-Salpêtrière are rather satisfied with the currently used PPR (it must be noted that the local Hospital Information System (HIS) offers nearly no medical services today). Compared to an unstructured "bag" that contains everything about a patient medical history, it is flexible enough for their daily practice. From a writer's point of view, it allows freedom of expression concerning what to record, to which level of detail, and how to describe it. However, from the reader's point of view, there is no consensus about the browsing facility through the PPR, when available. The lack of organization is considered either as a weak point for data retrieval under time pressure, or as a strong point for data mining in an exploratory way that respects individual habits and specific requirements of medical specialties¹⁰.

Nevertheless, information technology has changed the culture of the working environment dealing with patient care. Nowadays, clinicians are not alone operating with the patient and work-flow constraints should find a solution with the computerization of patient records. In this context, physicians have expressed the wish to have an overview of the complete medical record on one or two pages (screens) with the last conclusions as well as alerts that should not be missed in reading the record. Then when clicking on a word summarizing some episode of the medical record, it should be possible to retrieve all the documents describing that episode. This approach allows to combine the two opposite facets of the medical record, synthetic when clinicians need to get an overall idea and exhaustive according to legal requirements to display all the available information when developing a specific item.

A Mock-Up Demonstration

To refine the specification of the emerging system, we decided to develop a mock-up of a document-based EMR. The goal of this experiment was to evaluate more precisely the needs and wishes of potential users by confronting them with a running mock-up. It was also a concrete means to propose them new functionalities and to involve them in their definitions, or at least identification.

METHODS

Mock-Up Objectives

The objectives for this mock-up were twofold. On the users' side, we wanted to assess their acceptability of a future system for medical document management. On the project side, according to the state-of-the-art, we wanted to devise and evaluate the kind of Natural Language Processing (NLP) technologies that could be used, and at what cost, to provide an expected functionality.

Design Philosophy

The driving idea is to consider the EMR not as a collection of patient related medical data stored in a relational database that can be presented in a useful manner to the user, but rather as a collection of documents that primarily mimics the PPR. The "base documents" are strictly true to those health care professionals consult and produce in their daily practice. These include external documents related to the patient, as well as internal ones produced during the current hospitalization. The patient record appears then primarily as a collection of multimedia documents.

But, as pointed out previously, replicating the PPR in an EMR is not sufficient. Additional facilities must be added to enable other access modes to documents than just database queries. This problem pertains to the underlying hypertextual model to be designed to facilitate browsing through the medical record in accordance with clinicians' needs.

An important point was to suggest synthetic views of the various documents the actual PPR consists of. The first role of "synthesis documents" is to allow, as *documents*, a quick acknowledgment of the medical contents of the record. Their second role is to provide, as *hyper-documents*, an immediate and relevant access to this contents in an exploration mode suited to current practices in hospital departments, *i.e.* according to the medical specialty, and even to the department. Such a requirement implies that several access paths across the EMR are available.

Implementation

Mock-up development was based on WWW philosophy and available tools. On the server side, each document corresponds to an "HTML page" that is either

served or generated on the fly.² The EMR can therefore be browsed through any HTML 2.0 dedicated available browser, which constitutes the client user interface.

The mock-up presents an actual patient record in rheumatology for consultation only. A great part of the original paper documents have been keyboarded, consistently anonymized (dummy names, locations and dates have been used), and marked up. They are presented in a form as close as possible to their paper version. The mock-up contains more than one hundred base documents (text reports, images, raw data) covering 8 years of care of a single patient.

Evaluation And Refinement Cycle

Once a first version was implemented, the mock-up was demonstrated to hospital physicians, individually or by small groups. These physicians were involved in the project as members of Medical Advisory Boards (MABs), at national and international levels³. Their reactions, suggestions, and criticisms were reported in order to refine their needs and wishes about such a system and its functionalities. Then the mock-up entered a cyclic development phase where the remarks were, when possible, taken into account for a new presentation. The current version is the fourth and illustrates locally most of the browsing functionalities expected by potential users.

RESULTS

Base Documents

Base documents from the PPR were categorized according to medical criteria: on one side the Patient Discharge Summaries (PDSs) from previous hospital stays, on the other side the examination results where imaging and paraclinical data are distinguished.

Whatever the originating department, every PDS "almost" follows the same plan. Some information is invariant (originator: author/department/hospital, destination, date, etc.), and the rest is made of semi-structured text with typical sections containing arbitrary free text paragraphs (reason for admission, antecedents, history, conclusion, discharge treatment, etc.). This structure is explicitly encoded in the text with some specific added SGML mark-up¹³ among HTML tags.

Lab results were categorized according to the kind of test (blood gas, bacteriology, urinary tests, etc.), imagery data according to the investigation mode (invasive, radiology, echography, scintigraphy, etc.). Each document contains some test identifier in addition to the same invariant information as for PDSs. Raw data

are essentially presented in tables, but additional textual reports may be attached.

Synthesis Documents

Synthesis documents are expected to provide the reader with a fast and pertinent account of the whole PPR, of a single document, or of some set of documents. Most syntheses are up to now virtual documents. They do not have their paper counterparts as full documents, though the information they provide is dispatched in existing base documents or may be reconstructed.

Syntheses are either synopses, that give an overview of a whole, or summaries, that only report important information. Both may be either derived from the structure of documents, which also reflects their informational contents (*e.g.* table of contents, document author), or composed by physicians (*e.g.* problem list, imaging reports). In every case, synthesis documents constitute different ways to access more specific documents.

Document summary. Single document syntheses are used mostly for textual documents (PDSs, imagery reports) for which they provide a summary. They are based on their internal tagged structure and are automatically generated. Relevant information is presented without details. Basically, the summary reports the source (date/hospital/department), the "reason for admission" section, the "conclusion" section, the author, and the plan of the original document based on its headings.

Synopses. Synopses can also be produced for several documents from which they aggregate information. The first kind of useful syntheses that give an overview on some parts of the patient record are catalogs of documents sharing common properties. The first catalog is the complete chronological list of base documents which is lacking in the PPR, and where each document is referred to by its date, type, origin. Other catalogs may be generated from various user criteria. Raw data issued from multiple similar tests can be aggregated over a period of time to provide a synthetic view on the evolution of clinical parameters. Graphical accounts of such data can be presented in connection with the nursing record.

Global syntheses. The whole patient record syntheses are supposed to provide the reader with immediate critical medical information. Such synthesis documents are typically based on the informational contents of all the other documents in the patient record, but not on their structure (though it could help). These documents in a fully automated EMR could be extracted by contents analysis methods applied to existing base documents, and/or elaborated by physicians.

The "medical card", as the first document presented to the user, constitutes the first page of the EMR. It

²The server we used is NCSA httpd1.4, with dedicated Unix (SunOS) scripts using sh, Perl, wais.

³There were Belgian, Swiss, English, and French MABs.

contains a short list of significant medical information about the patient: (i) alerts that should not be missed; (ii) reason for current admission; (iii) main known diagnosis; (iv) conclusion extracted from the last PDS; (v) a time-oriented synoptic graph of previous hospitalizations from which the corresponding PDSs can be accessed; and (vi) pointers to other relevant synthesis documents.

Another important medical document is the “**synthesis of current hospitalization**”. It contains mostly full text and is structured like a PDS according to the following plan: (1) Reason for admission, (2) Antecedents, (3) Life habits, (4) Disease history, (5) Admission, (6) Evolution and tests, (7) Conclusion. This synthesis is the dynamically evolving draft version of the PDS to be written at discharge.

The “**antecedents list**” collects all past pathological events based on standardized medical terms. Several listing criteria are proposed: by type (medical or surgical) or by chronology. The “**disease history**” page, different by its contents, adheres to the same principles.

Navigating Through The EMR

Seeking information among huge amounts of documents is an important issue. The mock-up offers three access modes to the EMR documents that constitute its hypertext model. Although every access is obtained in the same way (*i.e.* by a mouse click), the underlying motivations and methods are quite different and do not serve the same information retrieval strategy.

Access to documents can be done according to their structural properties (typically their category) or their semantic properties, *i.e.* their contents. Hyperlinks relate some part of a document to another (or others) part(s). “Structural links” reflect the structural organization of the patient record. “Semantic links” relate some portion of a document (a medical term in context) with another document (or a part of it) according to a meaningful relation. A third access method is provided by arbitrary full text search.

Structural access. Structural access relies on the invariant structural properties of documents. Basically, catalogs provides comprehensive lists of various documents according to their type (PDS, radiology reports), at different levels of genericity (*e.g.* documents>test results>blood gas results) or according to some characteristics (*e.g.* same originating hospital). As for single textual documents, the table of contents reflects their structure and provides access to sections.

Hypertextual access. Hyperlinks relate some textual portion of a document with another document or a part of a document. Such links are based upon semantic relations in the medical context that are not systemat-

ically reflected by the document structure. Typically, an expression in a text, considered as an implicit reference either to another document (*e.g.* “*Coronarography practiced on Jan 15.*”) or to a clinical episode (*e.g.* “*Septicemia*”), is the origin of a link. The corresponding link destination is expected to deliver more detailed information; it is either a whole document (*e.g.* the corresponding coronarography report) or a selected part (*e.g.* the paragraph describing the septicemia episode in the appropriate PDS).

In the mock-up, hyperlinks have been set-up in most of the synthesized textual documents with the help of clinicians. Expressions using relevant medical vocabulary (*e.g.* the antecedents list, the synthesis of current hospitalization, alerts) are link origins enabling a direct access to original information, thus providing finer grained information and context.

Full text search. Full text retrieval capabilities offer a great flexibility in information retrieval. They have been implemented with freeWAIS for all textual documents related to a unique patient. Any document can be retrieved according to its “contents”. An arbitrary request over the EMR can be composed using words. The answer is the list of hits, *i.e.* the catalog of documents that contain the requested words.

DISCUSSION AND RELATED WORK

The proposal of an EMR as a Medical Document Management System, demonstrated by our mock-up, raised a strong interest among physicians. The approach accounts for current clinicians’ practice and does not enforce an *a priori* constraining model. It includes the PPR contents but is not limited to it. Similar experiments have been reported^{12,14}. The importance of EMR contents syntheses has been stressed, as summaries and especially as synoptic views¹⁵ or “time lines”¹⁶. Synthesis contents and presentations must be suited to the task, to the specialty¹⁷, or to the reading attitude and physician goals¹⁰. In any case, connections to original data must always be preserved.

Automating the production of synthesis documents and of the underlying hypertext structure in context becomes an issue. NLP should provide tools to facilitate this process. This necessitates that patient record models, but also social consideration on medical communications, *i.e.* work-flow models, are taken into account^{9,8}. NLP should also enhance document indexing and consequently allow various retrieval strategies. Information retrieval based solely on words is not sufficient; normalized terminologies, and even conceptual models, would be useful for this issue.

On the technical side, the generic WWW technology has proved its capability to enable a fast and easy development of a networked information server. It also

provided sufficient control on presentation and user interaction. However, as to the representation (and storage) of textual documents, strict HTML appears to be limited for it focuses mostly on form. The design of a hypertextual EMR requires to make the most of the medical contents. Documentary technology offers now better means to deal with contents. The use of a standard like SGML¹³ would permit the elaboration of medical document type definitions accounting for their medical informational contents. With the HyTime standard¹⁸, which allows for external links, the delivered hypertext structure may then be different for distinct tasks or users. Moreover, relying on international standards provides portability and independence.

However, if the importance of EMR consultation has been stressed, other critical issues exist. Authoring, editing, and annotating capabilities are necessary conditions for the elaboration of such an EMR. If WWW technology may be used as a major component over HISs for EMR consultation, it does not seem, in our opinion, to be sufficient for the other "authoring" tasks. Such tasks require medical knowledge about the composition of existing documents and of those which are edited. In this context, a project for a prototype of a health care professional "reader's workstation" as part of the EMR is planned in 3 departments of La Pitié-Salpêtrière.

CONCLUSION

The hypertextual documentary approach to the EMR, illustrated by a WWW mock-up, seems flexible enough to be accepted by different users and still useful because it does not impose an *a priori* patient record model and sticks to current hospital practice. On the users' side, it combines the advantages of free text with the benefits of computerization. However, if such an EMR is attractive for consultation, its setting-up raises authoring and processing issues that are not yet solved.

Acknowledgments

The authors would like to thank Pr. Bourgeois for providing medical material, their colleagues for fruitful discussions, and national and international MAB members for their collaboration and suggestions, especially Dr. Similowski, Dr. Zelter, and Dr. Capelle.

References

- Schatz BR and Hardin JB. NCSA Mosaic and the World Wide Web: Global hypermedia protocols for the internet. *Science* 1994;265.
- Buhle EL, Goldwein JW, and Benjamin I. OncoLink: A multimedia oncology information resource on the internet. In: SCAMC¹⁹, 1994:103-7.
- Kruper JA, Lavenant MG, Maskay MH, and Jones TM. Building internet accessible medical education software using the World Wide Web. In: SCAMC¹⁹, 1994:32-6.
- Fowler J, Kouramajian V, Maram S, and Devadhar V. Automated MeSH indexing of the World Wide Web. In: SCAMC²⁰, 1995:893-7.
- Cimino JJ, Socratous SA, and Clayton PD. Internet as clinical information system: Application development using the world wide web. *Journal of the American Medical Informatics Association* 1995;2(5):273-84.
- Tachinardi U, Furuie S, Bertozzo N, et al. Hypermedia patient data retrieval and presentation through WWW. In: SCAMC²⁰, 1995:551-5.
- Moidu K, Falsome JJ, and Nair S. Computer-based patient record: The essential data set approach. In: SCAMC¹⁹, 1994:462-6.
- van Ginneken AM. The structure of data in medical records. In: Bommel JHv and McCray AT, eds, *Yearbook of Medical Informatics 95 - The Computer-Based Record*. Schattauer, 1995:61-70.
- Frisse ME, Schnase JL, and Metcalfe E. Models for patient records. *Acad. Med.* 1994;69:546-50.
- Nygren E and Henriksson P. Reading the medical record. I. Analysis of physicians' ways of reading the medical record. *Computer Methods and Programs in Biomedicine* 1992;39:1-2.
- Tange HJ. The paper-based patient record: Is it really so bad? *Computer Methods and Programs in Biomedicine* 1995;48:127-31.
- Nygren E, Johnson M, and Henriksson P. Reading the medical record. II. Design of a human-computer interface for basic reading of computerized medical records. *Computer Methods and Programs in Biomedicine* 1992;39:13-25.
- Goldfarb CF. *The SGML handbook*. Oxford University Press, 1990.
- Adelhard K, Eckel R, Hölzel D, and Tretter W. A prototype of a computerized patient record. *Computer Methods and Programs in Biomedicine* 1995;48:115-9.
- Powsner SM and Tufte ER. Graphical summary of patient status. *Lancet* 1994;344:386-9.
- Plaisant C and Rose A. Exploring LifeLines to visualize patient records. Technical report CS-TR-3620, HCIL, University of Maryland, March 1996.
- van Ginneken AM and Stam H. Can one patient record accommodate the diversity of specialized care. In: SCAMC²⁰, 1995:406-10.
- DeRose SJ and Durand DG. *Making Hypermedia Work - A User's Guide to HyTime*. Kluwer Academic Publishers, 1994.
- Proceedings of the 18th Annual SCAMC*, Washington, 1994. Mc Graw Hill.
- Proceedings of the 19th Annual SCAMC*, New Orleans, 1995. JAMIA.